图书情报工作
第 64 卷 第 22 期 2020 年 11 月

数据管理与数据治理的比较分析及其对制定科学数据开放共享政策的启示*

盛小平 宋大成
上海大学图书情报档案系 上海 200444

摘 要:[目的/意义]比较分析数据管理与数据治理差异与联系,为制定科学数据开放共享政策提供参考。[方法/过程]运用比较分析法,解析数据管理与数据治理在定义与内涵、功能、目标、原则、焦点领域5个方面的异同,由此解析其对制定我国科学数据开放共享政策的启示。[结果/结论]数据管理与数据治理在定义与内涵、功能、目标、原则、焦点领域上都有显著差异,但两者也有内在联系。数据治理是成功实施数据管理的关键。认清两者的关系有助于明晰目前我国科学数据管理政策的不足之处,为今后完善科学数据管理办法提供参考,从而规划与制定实用的科学数据开放共享细则。

关键词:数据管理 数据治理 比较分析 科学数据 开放共享政策
分类号:G203
DOI:10.13266/j.issn.0252-3116.2020.22.001

1 引言

数据对于几乎所有组织(无论大小)的成功都非常重要,许多组织都试图实施数据治理计划[1]。数据治理也是我国实施国家大数据战略、加快建设数字中国的关键领域与重要抓手[2]。不过,数据治理与数据管理是一组容易混淆的概念,如今存在如下四种不同观点:①数据治理与数据管理是两种不同层次的活动,比如,数据治理侧重于从宏观上评估、指导和监督数据管理活动的执行,数据管理则侧重于从微观上采取恰当的行动来实现数据治理的决策[3];数据治理凌驾于数据管理之上,是对数据内容和权利的重新分配[4]。②数据管理包含数据治理,比如,数据治理是数据管理的一部分[5];数据治理是数据管理的中心活动[6];数据治理是公司数据管理策略的重要组成部分[7]。③数据治理包含数据管理,比如,数据治理包括数据服务、数据管理和数据协调3层含义[8]。④数据管理等同于数据治理,比如,数据治理是一种数据管理概念,涉及使组织能够确保在数据的整个生命周期中存在高数据质量的能力[9]。由此看来,人们对数据管理与数据治理的认识比较混乱与模糊。此外,国内有些政策文件如《大数据产业发展规划(2016-2020年)》,虽然强调了数据治理的重要性,但是没有提及数据管理;而另一些政策文件如《科学数据管理办法》,虽然强调了数据管理,但是忽视了数据治理。这不得不令人反思,数据治理与数据管理到底有何差异?我们如何有效利用数据管理与数据治理来实施我国科学数据开放共享?对于上述问题,鲜见相关文献进行深入探讨。因此,在这种情况下,很有必要厘清数据治理与数据管理的两者关系,解析两者的异同,以便制定有效的科学数据开放共享政策,并促进科学数据开放共享活动的实施。

2 数据管理与数据治理的比较分析

可以从如下5个方面比较分析数据管理与数据治理的异同。

2.1 两者定义与内涵的比较

简单地说,数据管理就是对数据、信息进行管理,以一种适合于通信、解释或处理的形式化的方式表示出来[10]。这是一种狭义的定义。从广义上讲,数据管理是对数据计划、政策、程序和实践的开发、执行和监

* 本文系国家社会科学基金项目"开放科学环境下的科学数据开放共享机制与对策研究"(项目编号:18ATQ007)研究成果之一。
作者简介:盛小平(ORCID:0000-0002-6341-6973),教授,博士,博士生导师,E-mail:shengxp68@126.com;宋大成(ORCID:0000-0003-3367-6915),硕士研究生。
收稿日期:2020-06-09 修回日期:2020-07-21 本文起止页码:4-10 责任编辑:杜杏叶

督,以便在整个生命周期内传递、控制、保护和提高数据和信息资产的价值[6]17。广义的数据管理包含数据治理。不过,数据管理通常关注数据元素的定义、如何存储、结构化和移动,以确保数据是可靠的和可利用的[11],是对数据生命周期全过程的管理。本文仅对狭义的数据管理与数据治理进行比较分析。

数据治理是对数据资产管理行使权力和控制(包括计划、监视和实施)[6]67。换句话说,数据治理是制定标准、规则、政策和实施监督以确保遵守数据管理最佳做法的过程[12]。它需要对组织(如企业)中数据可用性、相关性、使用性、完整性和安全性进行全面管理[13]。数据治理意味着在组织内[14]:①制定和管理数据的公共政策和事务,包括制定与数据有关的法律和政策以及数据质量控制和管理战略;②行使数据的决策权力,包括决定数据的含义、数据在哪里使用、数据的精确度以及需要遵循的规则;③控制数据的速度或大小,包括根据数据分类规则(机密、敏感、公开)共享数据;④控制数据行动,包括建立和监视适当的过程以管理组织各级数据的定义、生产和使用;⑤使数据得到控制,包括对数据访问进行分级管理、保护和审计;⑥对数据施加影响,包括使正确的人在正确的时间做出正确的数据决定。

2.2 两者功能的比较

数据管理与数据治理在功能上是不同的。一般说来,数据管理功能主要包括[15]:①数据架构管理:定义数据架构与数据资产管理蓝图。②数据开发:进行数据的分析、设计、实施、测试、部署、维护,将结构化和非结构化数据识别为有价值的组织资源。③数据操作管理:提供从数据采集到清除的操作支持。④数据安全管理:对数据隐私、机密与安全提供保护,提供合适的访问。⑤数据质量管理:定义、监测和改善数据质量。⑥参考数据和主数据管理:对组织各类主数据和参考数据提供管理。⑦数据仓库和商业情报管理:通过数据仓库或数据知识库提供数据存储与商业情报。⑧文件和内容管理:对数据文件与内容进行管理。⑨元数据管理:集成、控制和提供元数据。

有效的数据治理具有如下许多重要功能[16-17]:①支持组织的战略措施,帮助组织明确降低成本、改善决策流程、提高效益与效率的目标。②有助于开发和实施数据质量管理规程、指南和路线图,为组织提供独特的数据和信息管理能力。③有助于提供一致的定义、方法与程序,更好地获取数据和信息,提高数据质量、数据的安全性和一致性以及组织对数据的信任,减少整个组织中重复数据的存储以及因数据不准确而造成的风险。④有助于形成和应用标准化的数据质量流程与明确定义的质量指标,测量和跟踪整个组织的数据质量,改善不同业务部门之间的协调,帮助组织更有效地管理数据。⑤有助于方便和及时地访问组织中跨业务部门的数据和组织以外的业务数据,也有助于业务部门更好地为与组织中的其他部门共享信息做好准备。⑥促进组织内部网络的发展,以便在整个组织内有效地转移、交换和分配数据和其他资源,由此获得竞争优势和增强业务部门之间的互惠性、连通性和凝聚力。⑦促进组织数据质量文化建设,促进人们更广泛了解产品和业务部门之间的数据,更广泛地思考质量问题。⑧有助于人们遵守组织规则,增强对组织数据的问责,改进数据控制,防止因意外披露数据和信息而对业务产生负面影响。⑨改善与客户(数据提供者和数据用户)的关系和交流。⑩提高组织在行业与社会中的地位。

数据管理与数据治理虽然在功能上有明显的不同,但两者具有内在的关联性。这是因为两者都以数据为管理对象;数据管理功能聚焦于数据生命周期的管理,包括数据管理计划、记录数据、组织数据、改进分析程序、正确地保护敏感数据、在项目期间有足够的存储和备份、在项目结束后照顾数据、有效地共享数据以及寻找数据以便在新项目中重用,如此广泛的实践可以为数据治理奠定基础[18];数据治理聚焦于数据战略、质量、安全、合规与问责,是对数据内容和权利的重新分配[4],也是对数据管理的高层规划和控制[11]841,可以对如何执行数据管理功能起到指导[15]37、决策和监督作用[19]。

2.3 两者目标的比较

数据管理与数据治理在目标上并不完全相同。数据管理的目标是[6]18:①理解组织和其他利益相关者的数据与信息需求。②获取、存储、保护和确认数据资产的完整性。③确保数据和信息的质量,包括数据的准确性、完整性、集成性、及时性、相关性和有用性以及数据定义的明确性和共识。④确保利益相关者数据的隐私和机密性。⑤防止未经授权和不适当地使用数据和信息。⑥确保可以有效地使用数据来为组织增加价值。总之,组织数据管理的目标应该是具体的、可衡量的、可实现的(或可执行的)、现实的,就指定的目标时间范围来说是及时的[15]19。

不同于数据管理目标,数据治理的总目标是保障组织能够将数据作为资产进行有效管理[6]71。这个总

目标可细分为[20]:①促成更好的决策。②减少业务摩擦。③保护数据利益相关者的需求。④培训管理人员和工作人员采用共同的方法处理数据问题。⑤建立标准的、可重复的流程。⑥确保流程的透明度。⑦通过协调努力降低成本并提高效率。

由此看来,数据管理与数据治理两者在确保利益相关者数据需求、有效使用数据来提高效率或创造价值方面,目标是一致的。然而,数据管理目标的重点是了解机构或用户数据需求,组织并提供相关数据,使数据价值得到充分利用;而数据治理的重点是更有效管理数据资产,将数据转化为可操作的知识,并降低数据风险[21]。

2.4 两者原则的比较

拥有数据并不等于知道如何处理数据。数据是一种具有独特属性的资产,数据在使用时不像金融和实物资产一样会被消耗,需要从组织视角来看待数据管理。数据管理原则主要包括[6]21-23:①数据价值量化原则:数据的价值可以而且应该用经济术语来表示,因为数据是组织的一种有价值的资产。组织应该利用相关技术和制定一致的方法来测量数据的定性和定量价值,包括衡量低质量数据的成本和高质量数据的好处,以便利用数据做出更好的决策。②数据质量原则:组织必须理解利益相关者对数据质量的要求,并根据这些需求测量数据质量。③元数据管理原则:管理任何资产都需要有关该资产的数据(雇员人数、会计代码等)。那些用于管理和使用数据的数据称为元数据。元数据来源于一系列与数据创建、处理和使用有关的的过程,包括架构、建模、管理、治理、系统开发、信息技术、业务操作和分析。④计划管理原则。管理数据需要计划。即使是小型组织也可能拥有复杂的技术和业务流程事务。数据创建于不同的地方(如不同部门或岗位、个人),且在这些地方之间移动以供使用。为了协调工作和保持最终结果一致,需要从架构和工作流程角度对数据进行计划与管理。⑤团队协作原则:数据管理需要一系列的技能和专业知识,单个团队无法管理组织的所有数据,需要跨职能部门与人员的协作。⑥动态管理原则:数据是流动的和变化的,需要用动态发展的观点看待数据管理。⑦生命周期管理原则:不同类型的数据具有不同的生命周期特征,它们有不同的管理需求,组织需要进行数据的生命周期管理。⑧风险管理原则:管理数据包括管理与数据相关的风险。数据除了是一种资产外,还代表组织的风险。因为数据可能丢失、被盗或误用。各组织必须考虑其使用数据的道德影响,并把数据相关风险作为数据生命周期的一部分进行管理。⑨技术关联性原则:数据管理与信息技术密切相关。管理数据需要一种可以确保技术服务于组织战略数据需求的方法。⑩领导支持原则:有效的数据管理需要领导的承诺与支持。

不同于数据管理原则,数据治理原则是[20]:①一致性:确保数据治理与业务需求保持一致,确保对数据的共同理解。②多层次:数据治理发生在整个组织和各部门之间。③诚实性:数据治理参与者在彼此的交往中、相关决策中是诚实和坦率的。④透明性:数据治理和管理工作流程将是透明的。⑤可审核性:数据治理决策、流程和控制是可审计的。⑥问责与制衡:将以一种制衡方式明确个人和团队在数据管理活动中的责任。⑦标准化:数据治理将引入和支持数据的标准化。⑧变革管理:数据治理支持主动和被动的变革管理活动。

2.5 两者焦点领域及相关活动的比较

数据管理与数据治理分别拥有各自的焦点领域及其不同的主要活动。数据管理的焦点领域主要是数据架构(data architecture)管理、数据质量管理、数据仓库和商业情报管理、数据存储和操作管理、数据集成和互操作管理、文件和内容管理、参考和主数据管理、数据建模和设计管理、元数据管理、数据安全管理等[6]45-46。

一般说来,数据治理的焦点领域包括如下7个方面[15]16-17,[22-23]:①以数据管理计划为重点的数据治理;②以数据结构为重点的数据治理;③以数据质量为重点的数据治理;④以数据仓库和商业情报为重点的数据治理;⑤以数据访问为重点的数据治理;⑥以数据使用为重点的数据治理;⑦以隐私、法规、安全为重点的数据治理;⑧以数据管理控制为重点的数据治理。

基于上述数据管理与数据治理焦点领域的划分,可以进一步比较分析其相关活动的异同。

(1)不同焦点领域及相关活动的比较。数据管理与数据治理的大多数焦点领域及其相关活动上是不同的(见表1)。

(2)相同焦点领域及相关活动的比较。虽然数据管理与数据治理有3个相同的焦点领域——数据质量管理、数据架构管理、数据仓库和商业情报管理,但是各自强调的相关活动的侧重点不同(见表2)。其中,在数据质量管理领域,数据管理主要是围绕如何定义、执行、改进、监测数据质量而开展活动的,而数据治理是围绕数据的质量、完整性或可用性而开展活动的。

盛小平，宋大成.数据管理与数据治理的比较分析及其对制定科学数据开放共享政策的启示[J].图书情报工作,2020,64(22):4－10.

表 1　不同的焦点领域及其活动列表

数据管理		数据治理	
焦点领域	相关主要活动	焦点领域	相关主要活动
数据存储和操作管理	定义数据存储要求;识别数据库使用方式;定义数据访问要求;备份与恢复数据;开发数据库实例;管理数据库性能;管理测试数据集;管理数据迁移;了解数据库技术特性;管理、监控与评估数据库技术[6]194－208	以数据管理计划为重点的数据治理	界定本组织的数据管理;开展数据治理准备状态评估;制定数据治理战略;定义数据治理操作框架;制定数据治理目标、原则和政策[6]79－83;建立数据专业人员角色和组织;任命数据专员;建立数据治理和管理组织;计划和赞助数据管理项目和服务;评估数据资产价值和相关成本[15]16
数据集成和互操作管理	定义数据集成与生命周期要求;实施数据发现;记录数据沿袭;进行数据配置;收集业务规则;设计数据集成结构;对数据集成、接口、通讯和数据服务进行建模;将数据源映射到目标;设计数据业务流程;开发数据服务、数据流与数据迁移方法;制定数据出版方法;维护数据集成和元数据;实施与监控数据集成和互操作[6]286－293	以数据管理控制为重点的数据治理	监管数据专业机构和人员;协调数据治理活动;管理和解决数据相关问题;监控并确保法规遵从;倡议数据标准和程序;监控和执行数据政策、标准和架构的一致性;监督数据管理项目和服务;测评数据及其相关工作的价值;交流和提升数据资产价值[15]16－17;从事数据管理项目;实施变革管理;从事问题管理[6]84－87
文件和内容管理	进行文件与记录管理计划;制定内容战略;制定内容处理政策;定义内容信息架构;获取记录与内容;管理版本和控制;管理保留和处置;审核文件或记录;提供访问、搜索和检索;交付文件和记录[6]323－329	以数据访问为重点的数据治理	确定可利用的数据和数据集;确保数据可被发现和使用;确保数据透明度;处理数据访问请求;建立公平的访问条件;尊重同意条款;管理可耗尽的资源;制定滥用数据的制裁措施;建立有效和协调的数据访问治理系统;提供数据访问的费用;确保存储库的可持续性[24]
参考和主数据管理	定义参考和主数据动力和需求;评估参考和主数据来源;定义参考和主数据架构方法;建模主数据和参考数据集;定义管理和维护流程;制定参考和主数据治理政策[6]371－375	以数据使用为重点的数据治理	描述数据的业务用途,确定数据在多大程度上是组织范围内的资产及其理想的使用行为;确定持续交流业务数据的机制;确定共享和重用数据的机会;监管环境如何影响业务数据的使用[24]
元数据管理	定义元数据策略;了解元数据需求;定义元数据架构;创建和维护元数据;查询、报告和分析元数据[6]434－440	以隐私、法规、安全为重点的数据治理	通过支持访问管理和安全要求,帮助保护敏感数据;协调框架和举措;帮助评估风险,定义管理风险的控制措施;帮助执行监管、合同、架构合规要求;识别利益相关者,建立决策权,明确责任[21]
数据建模和设计管理	制定数据建模计划;构建数据模型;审查数据模型;维护数据模型[6]152－159		
数据安全管理	了解数据安全需求;确定数据安全政策;定义数据安全标准[6]245－256		

在数据架构管理领域,数据管理主要是围绕建立数据架构框架和开展数据架构实践而开展活动的,而数据治理主要是围绕数据架构的政策、标准、元数据方案与权责管理而开展活动的。在数据仓库和商业情报管理领域,数据管理主要是围绕了解、定义、开发和维护数据仓库和商业情报而开展活动的,数据治理主要是围绕制定数据使用规则、明确利益相关者权责、确定数据资产价值而开展活动的。

表 2　相同的焦点领域及其活动列表

相同焦点领域	数据管理主要活动	数据治理主要活动
数据质量管理	定义高质量数据;定义数据质量战略;确定关键数据和业务规则;执行初始数据质量评估;确定和优先考虑潜在的改进;定义数据质量改进目标;管理数据质量原则;测量和监控数据质量;制定管理数据问题的操作程序;创建数据质量服务等级协议;形成数据质量报告[6]473－484	为数据质量设定方向;监测数据质量;报告质量焦点举措状况;确定数据质量利益相关者及其责任[21];确定数据准确性、及时性、完整性和可靠性方面的质量标准;确定数据质量评估方法与计划[23]
数据架构管理	开发和维护组织数据模型;定义和维护数据技术架构[15]17;评估现有的数据架构规范;制定组织数据架构路线图;管理组织项目需求;与企业架构集成[6]110－115	确保一致的数据定义;支持数据架构政策和标准;支持元数据方案、面向服务架构、主数据管理、企业数据管理;识别利益相关者,建立决策权,明确责任[21]
数据仓库与商业情报管理	了解数据仓库和商业情报要求;定义和维护数据仓库和商业情报架构;开发数据仓库和数据集市;填充数据仓库;实现商业情报组合;维护数据产品[6]394－401	制定数据使用和数据定义规则;识别利益相关者,建立决策权,明确责任;确定软件开发生命周期嵌入的治理步骤和项目循环;澄清数据资产和数据相关项目的价值[21]

3　数据管理与数据治理的区别与联系

通过上述比较分析,笔者发现,数据治理与数据管理是两个不同的概念,两者在定义与内涵、功能、目标、原则、焦点领域等方面都有明显差异(见表3)。

总之,数据管理与数据治理的区别主要体现在:①数据治理是对数据管理的规划、决策、监督和控制,而数据管理是对数据生命周期活动(包括数据的规划、收集、描述、综合、分析、保存、评价、再利用)的管理。②数据管理的主要目标是了解用户数据需求,组织并

表3 数据管理与数据治理的区别

比较维度	数据管理	数据治理
定义与内涵	对数据计划、政策、程序和实践的开发、执行和监督	制定标准、规则、政策和实施监督以确保开展数据管理
功能	聚焦于数据生命周期的管理	聚焦于数据内容和权利的重新分配
目标	重点是了解机构或用户数据需求	重点是更有效管理数据资产
原则	侧重于对数据属性、过程、技能的管理要求	侧重于对数据责任、质量、标准、合规的管理要求
焦点领域	数据质量、数据架构、数据仓库和商业情报、数据存储和操作、数据集成和互操作、文件和内容管理、参考和主数据管理、数据建模和设计、元数据、数据安全管理	数据治理政策、标准、策略，数据架构，数据质量，数据仓库和商业情报，数据访问，数据隐私、法规与安全，数据原则，管理支持

提供高质量的相关数据，使数据价值得到充分利用；而数据治理的主要目标是确保数据资产能够得到有效管理与运营，实现其价值的最大化和降低各种数据风险。③数据管理主要遵循数据价值量化原则、数据质量原则、元数据管理原则、数据规划原则、动态管理原则、生命周期管理原则、技术关联性原则、领导支持原则，侧重于对数据属性、过程、技能的管理要求；而数据治理更加强调按照一致性、透明性、合规性、可审计、标准化、问责制原则来管理数据，侧重于对数据责任、质量、标准、合规的管理要求。④数据治理与数据管理都关注数据架构管理、数据质量管理、数据仓库和商业情报管理3个焦点领域，但其相关活动上是不同的。其他数据管理焦点领域包括数据存储和操作、数据集成和互操作、文件和内容管理、参考和主数据管理、数据建模和设计、元数据、数据安全管理等。其他数据治理焦点领域包括数据管理计划、数据管理监控、数据访问、数据使用、数据隐私和法规以及安全问题等。⑤数据治理并不涉及数据管理的点点滴滴，而是主要关注数据的可用性、完整性、一致性、可靠性、安全性、架构、质量、隐私、合规[25]。

另一方面，数据治理与数据管理具有内在的紧密联系。这主要体现在如下几方面：①数据治理在本质上属于广义的数据管理范畴[26]，数据治理是成功实施狭义的数据管理的关键。②对于某个组织来说，数据管理和数据治理都是缺一不可的。若没有有效的数据管理，组织将会面临许多问题，包括[10]58-59：组织内的信息系统不能相互连接，因为它们没有被定义和开发来一起工作，信息系统只能在最基本的技术层面上被连接；数据不能在信息系统之间共享，即使从技术上连接信息系统是可能的，但由于在不同的信息系统中使用不兼容的数据定义，在这些系统之间共享数据通常仍然是不可能的；沟通中断，信息丢失；不必要地转录和重新键入数据；重复信息和数据需求分析工作，增加研发时间和成本；组织的竞争优势被削弱；引发用户和员工的挫折感。若缺少数据治理，组织就会不知道如何有效处理数据、跟踪与监督数据利用情况，也不知道哪些数据是可信的、安全的、兼容的、保密的和可访问的，更不能确保组织能够最大限度地利用数据治理投资和尽可能地将数据破坏的风险降到最低。③数据管理和数据治理在确保数据质量与安全、最大限度地挖掘与利用数据资产价值等方面具有相同的目标。④数据管理和数据治理都强调领导支持、风险管理、协同管理、动态管理的管理原则。

4 对制定我国科学数据开放共享政策的启示

认清数据治理与数据管理的区别与联系，有助于审视国内外科学数据管理政策，并指导科学数据开放共享实践，特别是对于制定我国科学数据开放共享政策而言，可提供如下两点启示。

第一，可以明晰科学数据管理政策的不足之处，为今后完善科学数据管理办法提供参考。2018年3月17日国务院办公厅颁布的《科学数据管理办法》首次明确了国务院科学技术行政部门、国务院相关部门、省级人民政府相关部门、科研院所、高等院校和企业等法人单位、科学数据中心在科学数据管理过程(包括采集、汇交、保存、共享与利用)中以及在“保密与安全”环节上的职责与做法[27]，对于加强我国科学数据管理具有十分重要的促进作用。《科学数据管理办法》虽然确立了“开放为常态、不开放为例外”的一般性原则，但是并没有说明到底如何实施科学数据的开放共享。尤其是《科学数据管理办法》没有把科学数据治理条款纳入其中，既没有明确界定数据的归属权、数据开放的时限、数据的出版和引用机制等问题[28]，也没有规范科学数据共享中的个人数据保护、隐私治理、数据质量、数据权利、数据服务等关键问题。《科学数据管理办法》虽然强调了科学数据管理思想，但是忽视了科学数据治理措施。事实上，高效且成功的科学数据管理离不开科学数据治理的支持。成功实施全国范围

内的科学数据开放共享的前提条件之一就是需要建立健全科学数据开放共享治理体系。这包括建立"国家——地方——机构——个人"四级数据治理主体组成的治理结构体系,明确不同利益相关者在科学数据开放共享中的权利、义务与职责,按照可用性、完整性、一致性、可靠性、安全性、标准化、合规要求确立科学数据质量跟踪与评价机制、科学数据安全与隐私保护机制、科学数据开放共享机制。只有把科学数据管理与科学数据治理有效融合起来,并在今后《科学数据管理办法》修订过程中适当添加科学数据治理措施,才能保障科学数据开放共享得到更有效的实施。

第二,规划与制定实用的科学数据开放共享细则。在《科学数据管理办法》指导下,2019 年 2 月 11 日中国科学院率先在国内出台了行业性的科学数据管理细则——《中国科学院科学数据管理与开放共享办法(试行)》[29]。该试行办法较好地体现了综合运用数据管理与数据治理的思想来规范科学数据管理与开放共享行为。这包括:①把数据管理理念与数据治理思想结合起来,制定了"科学数据工作遵循统筹协调、规范管理、安全可控、持续发展的基本原则,落实主体责任,创新管理机制,加强能力建设,促进开放共享"的总体原则(见第一章《总则》第四条)。②通过明确中国科学院网络安全和信息化领导小组、中国科学院网络安全和信息化领导小组办公室、中国科学院院机关各部门、院属法人单位、中国科学院科学数据中心的职责,初步构建了中国科学院科学数据治理体系(见第二章《职责分工》),使科学数据管理工作落实到具体的单位或部门。③运用数据管理思想,规范了科技项目数据和论文关联数据的汇交与管理(见第三章、第四章),明确了院属机构或单位开放共享科学数据的做法(见第五章)。④确立了科学数据管理与开放共享的保障机制、安全保密机制(见第六章、第七章),既强化了科学数据管理思想,也灵活运用了科学数据治理措施。总之,该试行办法是中国科学院实施国家大数据战略的重要举措,对于促进科学数据开放共享具有开创性的历史意义。不过,美中不足的该办法没有全面、详细界定个人(如科研人员、用户)在科学数据管理与开放共享中的权利、责任与义务,致使并未构建完善的科学数据治理体系。此外,目前鲜见国内其他机构出台科学数据开放共享细则。这就意味着我国仍需加强科学数据管理与治理的政策研究与开放共享的实践研究,通过融入数据管理与数据治理思想,建立健全科学数据开放共享政策体系,以确保我国科学数据开放共享实践得到可持续发展。

参考文献:

[1] PLOTKIN D. Data stewardship: an actionable guide to effective data management and data governance[M]. Waltham: Morgan Kaufmann,2014:1.

[2] 李海韵. 习近平:实施国家大数据战略加快建设数字中国[EB/OL]. [2020-07-20]. http://www.xinhuanet.com/2017-12/09/c_1122084706.htm.

[3] 包冬梅,范颖捷,李鸣. 高校图书馆数据治理及其框架[J]. 图书情报工作,2015,59(18):134-141.

[4] 任亚忠. 从数据管理走向数据治理——大数据环境下图书馆职能的转变[J]. 四川图书馆学报,2017(4):6-9.

[5] HARPER J. Distinguishing data management from data governance[EB/OL]. [2020-07-20]. https://www.dataversity.net/distinguishing-data-management-from-data-governance/.

[6] DAMA International. Data management body of knowledge[M]. 2nd ed. Basking Ridge: Technics Publications, 2017:35.

[7] Data Republic. Data governance vs data management[EB/OL]. [2020-07-20]. https://www.datarepublic.com/blog/data-governance-vs-data-management.

[8] 林宁. 从数据管理到数据治理[N]. 中国海洋报,2015-07-08(003).

[9] 数据治理是一种数据管理概念——维基百科[EB/OL]. [2020-07-20]. http://www.esensoft.com/industry-news/data-governance-899.html.

[10] GORDON K. Principles of data management: facilitating information sharing[M]. 2nd edition. Swindon: BCS Learning and Development Ltd, 2013:57.

[11] AL-RUITHE M, BENKHELIFA E, HAMEED K. A systematic literature review of data governance and cloud data governance[J]. Personal and ubiquitous computing, 2019,23(5/6): 839-859.

[12] Enterprise Data Management Council. Data management capability assessment model (DCAM) [EB/OL]. [2020-07-20]. https://dgpo.org/wp-content/uploads/2016/06/EDMC_DCAM_-_WORKING_DRAFT_VERSION_0.7.pdf.

[13] IBM. What is data governance? [EB/OL]. [2020-07-20]. https://www.ibm.com/analytics/data-governance.

[14] SEINER R S. Non-Invasive data governance: the path of least resistance and greatest success [M]. Basking Ridge:Technics Publications, 2014:21-24.

[15] MOSLEY M, BRACKETT M, EARLEY S, et al. The DAMA guide to the data management body of knowledge (DAMA-DMBOK)[M]. Bradley Beach: Technics Publications, 2009: 19-20.

[16] BHANSALI N. The role of data governance in an organization [M]// BHANSALI N. Data governance : creating value from information assets. Boca Raton: CRC Press, 2014:1-18.

[17] ZHANG J. Operationalizing data quality through data governance [M]// BHANSALI N. Data governance : creating value from in-

formation assets. Boca Raton: CRC Press, 2014:65-92.

[18] BRINEY K. Data management for researchers: organize, maintain and share your data for research success[M]. Exeter: Pelagic Publishing, 2015:6.

[19] SPACEY J. Data governance vs data management[EB/OL]. [2020-07-20]. https://simplicable.com/new/data-governance-vs-data-management.

[20] The Data Governance Institute. Goals and principles for data governance[EB/OL]. [2020-07-20]. http://www.datagovernance.com/adg_data_governance_goals/.

[21] COHN B L. Data governance: a quality imperative in the era of big data, open data and beyond[J]. I/S: a journal of law and policy for the information society, 2015,10(3):811-826.

[22] THOMAS G. The DGI data governance framework[EB/OL]. [2020-07-20]. http://www.datagovernance.com/wp-content/uploads/2014/11/dgi_framework.pdf.

[23] The Expert Advisory Group on Data Access(EAGDA). Governance of data access[R/OL]. [2020-07-20]. https://wellcome.ac.uk/sites/default/files/governance-of-data-access-eagda-jun15.pdf.

[24] KHATRI V, BROWN C V. Designing data governance[J]. Communications of the ACM,2010, 53(1):148-152.

[25] MANDULA L. What is the difference between data governance and data Management[EB/OL]. [2020-07-20]. http://beckerassociates.ca/the-difference-between-governance-and-management/.

[26] 樊振佳. 企业科研数据治理与管护主体:NSR的实践与剖析[J]. 图书情报工作,2017,61(1): 56-63.

[27] 国务院办公厅. 国务院办公厅关于印发科学数据管理办法的通知[EB/OL]. [2020-07-20]. http://www.most.gov.cn/mostinfo/xinxifenlei/fgzc/gfxwj/gfxwj2018/201804/t20180404_139023.htm.

[28] 邢文明,洪程. 开放为常态,不开放为例外——解读《科学数据管理办法》中的科学数据共享与利用[J]. 图书馆论坛,2019,39(1):117-124.

[29] 中国科学院. 中国科学院关于印发《中国科学院科学数据管理与开放共享办法(试行)》的通知[EB/OL]. [2020-07-20]. http://www.cas.cn/tz/201902/P020190220358041915907.pdf.

作者贡献说明:
盛小平:论文撰写与修改;
宋大成:资料收集与内容分析。

A Comparative Analysis of Data Management and Data Governance and Its Enlightenment to the Formulation of Open Sharing Policies of Scientific Data

Sheng Xiaoping Song Dacheng
School of Library, Information and Archives, Shanghai University, Shanghai 200444

Abstract: [**Purpose/significance**] This paper compares and analyzes the differences and connections between data management and data governance, in order to provide references for the formulation of scientific data sharing policies. [**Method/process**] By using the comparative analysis method, this paper analyzed the similarities and differences between data management and data governance in five aspects: definition and meaning, functions, objectives, principles and focus areas, and then analyzed its enlightenment to the formulation of open sharing policies of scientific data in China. [**Result/conclusion**] There are significant differences between data management and data governance in terms of definition and meaning, functions, goals, principles and focus areas, but they are also intrinsically related. Data governance is the key to the successful implementation of data management. Understanding the relationship between them is helpful to clarify the deficiencies of current scientific data management policies in China, provide references for improving scientific data management methods in the future, and plan and formulate practical rules for the open sharing of scientific data.

Keywords: data management data governance comparative analysis scientific data open sharing policy